

LLM-in-the-loop: Leveraging Large Language Model for Thematic Analysis

Year: 2023 | Citations: 110 | Authors: Shih-Chieh Dai, Aiping Xiong, Lun-Wei Ku

Abstract

Thematic analysis (TA) has been widely used for analyzing qualitative data in many disciplines and fields. To ensure reliable analysis, the same piece of data is typically assigned to at least two human coders. Moreover, to produce meaningful and useful analysis, human coders develop and deepen their data interpretation and coding over multiple iterations, making TA labor-intensive and time-consuming. Recently the emerging field of large language models (LLMs) research has shown that LLMs have the potential replicate human-like behavior in various tasks: in particular, LLMs outperform crowd workers on text-annotation tasks, suggesting an opportunity to leverage LLMs on TA. We propose a human-LLM collaboration framework (i.e., LLM-in-the-loop) to conduct TA with in-context learning (ICL). This framework provides the prompt to frame discussions with a LLM (e.g., GPT-3.5) to generate the final codebook for TA. We demonstrate the utility of this framework using survey datasets on the aspects of the music listening experience and the usage of a password manager. Results of the two case studies show that the proposed framework yields similar coding quality to that of human coders but reduces TA's labor and time demands.